# Structure of the Toll-Spätzle complex, a molecular hub in *Drosophila* development and innate immunity

Christoph Parthier[a,1], Marco Stelter[a,1], Christian Ursel[a,1], Uwe Fandrich[a], Hauke Lilie[a], Constanze Breithaupt[a], and Milton T. Stubbs[a,b,2]

[a]Institut für Biochemie und Biotechnologie and [b]Mitteldeutsches Zentrum für Struktur und Dynamik von Proteinen, Martin-Luther-Universität Halle-Wittenberg, 06120 Halle (Saale), Germany

Edited by Jie-Oh Lee, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea, and accepted by the Editorial Board March 18, 2014 (received for review November 6, 2013)

***Drosophila*** **Toll receptors are involved in embryonic development and the immune response of adult flies. In both processes, the only known Toll receptor ligand is the human nerve growth factor-like cystine knot protein Spätzle. Here we present the crystal structure of a 1:1 (nonsignaling) complex of the full-length Toll receptor ectodomain (ECD) with the Spätzle cystine knot domain dimer. The ECD is divided into two leucine-rich repeat (LRR) domains, each of which is capped by cysteine-rich domains. Spätzle binds to the concave surface of the membrane-distal LRR domain, in contrast to the flanking ligand interactions observed for mammalian Toll-like receptors, with asymmetric contributions from each Spätzle protomer. The structure allows rationalization of existing genetic and biochemical data and provides a framework for targeting the immune systems of insects of economic importance, as well as a variety of invertebrate disease vectors.**

Toll signaling | embryonic morphogenesis | protein evolution

The model organism *Drosophila melanogaster* has yielded valuable insights into a number of fundamental biological processes. Genetic screens of flies subjected to chemical saturation mutagenesis have revealed a host of genes necessary for development, including *Toll* (from the German for "fantastic"), identified as a major determinant in the development of dorsal-ventral polarity (1–3). *Toll* encodes for a type I integral membrane protein with a large N-terminal extracellular domain consisting of a series of leucine-rich repeats (LRRs) (4) flanked by cysteine-rich motifs (5). The cytoplasmic intracellular C-terminal domain shares significant similarities with the mammalian interleukin-1 receptor (6, 7) and thus is termed the Toll-interleukin receptor (TIR) domain.

Microinjection experiments have identified the *spz* gene product Spätzle as the ligand for Toll (8, 9). The morphogen, present throughout the perivitelline space in the form of an inactive precursor proSpätzle, undergoes activation cleavage by the serine protease Easter (10) (itself activated during development by a spatially confined proteolytic cascade), leading to an extracellular gradient of activated Spätzle in the developing embryo (reviewed in ref. 11). Spätzle-mediated activation of the Toll receptor results in nuclear localization of the NF-κB/rel transcription factor Dorsal (12), eliciting transcription of further ventral-specific differentiation genes. In addition, the absence of nuclear Dorsal protein on the dorsal side of the embryo derepresses another distinct class of genes (13).

The identification of a role for Toll and Spätzle in the innate immune response of adult flies (14) led to increased interest in the Toll signaling pathway, particularly after the discovery that Toll-like receptors (TLRs) are involved in activation of the mammalian adaptive immune response (15, 16). The mammalian TLRs are activated directly by bacterial, fungal, or viral components known as pathogen-associated molecular patterns (17), and representative structures of seven ectodomains (ECDs) of the 11 known mammalian TLRs (TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, and TLR8) are available (18–21). Common to all of these receptors is ligand-induced activation via homodimer or heterodimer formation, although the specifics of this process can vary considerably.

Despite the pioneering role played by the *Drosophila* system, the Toll pathway is far less well characterized than its mammalian counterpart. Studies using chimeric EGF-Toll receptor reporter assays have demonstrated that Toll activation proceeds via receptor dimerization (22). Although nine *Toll*-related genes have been identified so far (5, 23), little is known about the functions of the other receptors. Toll-6 and Toll-7 have only very recently been identified as neurotrophin receptors in *Drosophila* (24), activated by the Spätzle paralogs Spz-2 and Spz-5. Toll signaling in insects differs substantially from that in mammals, in that the only known Toll ligand in both development and innate immunity is the Spätzle protein. In *Drosophila* innate immunity, the prohormone proSpätzle achieves its active form through maturation cleavage by the serine protease Spätzle-processing enzyme (25), analogously to its Easter-catalyzed activation in development (10).

As a result of alternative splicing, the proSpätzle precursor exists in various isoforms (26), most of which differ only in the length and sequences of their prodomains. Analysis of selected proSpätzle isoforms has revealed differences in their biophysical properties (27, 28). On Easter activation cleavage, the prodomains remain associated with the active cytokine in a noncovalent

## Significance

**During the earliest stages of fruit fly development, differentiation of the embryo into dorsal and ventral sections commences following the localized initiation of a proteolytic cascade that culminates in cleavage and activation of the human nerve growth factor-like cystine knot protein Spätzle. In turn, this activated ligand activates the Toll receptor, instigating an intracellular signal cascade that leads to location-specific cell differentiation. Both Toll and Spätzle are also integral to pathogen recognition in adult flies, where a similar proteolytic cascade results in triggering of the innate immune response. Despite functional similarities to the Toll-like receptors (TLRs) of mammalian innate immunity, the structure of the Toll-Spätzle complex described here exhibits a number of features that have not been observed in TLRs.**

Author contributions: M.T.S. designed research; C.P., M.S., C.U., U.F., and H.L. performed research; C.P., M.S., C.U., H.L., C.B., and M.T.S. analyzed data; and C.P., C.B., and M.T.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.-O.L. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates of the Toll-Spätzle complex have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 4LXR (deglycosylated) and 4LXS (glycosylated)].

[1]C.P., M.S., and C.U. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: stubbs@biochemtech.uni-halle.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1320678111/-/DCSupplemental.

complex termed Spätzle*, with a micromolar dissociation constant (28, 29). The mature Spätzle cystine knot domain itself is a disulfide-bonded covalent dimer related structurally to human nerve growth factor (NGF) (27, 30) that binds to Toll with low nanomolar affinity (31–33). The cystine knot domain molecule (also termed C106, the C-terminal 106 residue) takes the form of a T-shaped, covalently linked homodimer with a deep internal cavity (27), with N- and C-termini of each monomer in close proximity at the base of the cystine knot, and two prominent apical β-wings, each consisting of an intermolecular β-sheet incorporating residues $Gln^{75}$–$Phe^{94}$ of one protomer and residues $Tyr^{18}$–$Gln^{40}$ of the second protomer.

The stoichiometry of binding of Spätzle to Toll has been controversial. Both a one Spätzle dimer–one Toll receptor (1:1) complex and a one Spätzle dimer–two Toll receptor (1:2) complex have been reported (31, 32), as has a two Spätzle dimer–two Toll receptor (2:2) model (34), and preformed dimers of Toll in the absence of ligand have been described (32, 34). A model in which Spätzle binds to the N-terminal tip of Toll has been proposed based on a low-resolution cryoelectron microscopy (cryoEM) single-particle reconstruction (34), although this apparently has been ruled out by the recent characterization of a truncated Toll N-terminal domain (35). Here we report the crystal structure determination of the Spätzle dimer bound to the entire Toll receptor ECD in a 1:1 complex.

## Results

**Structure of the Toll ECD.** The ECD of Toll resembles a question mark, with a large semicircular N-terminal LRR domain ("ligand-binding" domain) and a second smaller ellipsoidal C-terminal LRR domain ("membrane-proximal" domain) (Fig. 1). As described recently (35), the N-terminal domain begins with a novel N-terminal capping region (amino acids 28–125; LRRNT1), composed of an open mixed β-sheet with a pronounced twist consisting of two β-hairpins followed by one parallel β-strand ($\downarrow_1\uparrow_2\downarrow_3\uparrow_4\uparrow_5$). Although the final secondary structure element of the LRRNT1 does not conform to any LRR motif, strand β5 extends the subsequent horseshoe-shaped β-sheet on the concave surface of the LRR domain, and thus we refer to this element as LRR0. The N-terminal LRR module (amino acids 126–553), which spans 182° of an asymmetric ellipse, consists of 17 conventional LRRs whose consensus sequence matches that seen in eukaryotic TLRs (36, 37). This module can be divided into three coplanar segments of differing curvature (LRRs 1–8, LRRs 9–12, and LRRs 13–15), giving rise to an asymmetric ellipse reminiscent of TLR4 (38).

Apart from LRR1/2 and LRR16/17, which link the LRR module to the neighboring capping structures, the LRRs exhibit nearly identical repeat lengths (11 LRRs are 24-aa long) with three- to five-residue β-strands forming the concave surface and a high degree of sequence similarity (*SI Appendix*, Fig. S1). The LRR module is terminated by a typical C-terminal capping motif (amino acids $Asn^{552}$–$Ser^{623}$; LRRCT1) (39), the first two β-strands of which add to and extend the parallel membrane distal LRR β-sheet, ending in an extended sequence stretch (amino acids $Leu^{616}$–$Ser^{623}$) containing the final C-cap $Cys^{618}$ that is disulfide-linked to $Cys^{567}$.

A second 171-residue LRR domain (the "dot" of the question mark, $Lys^{630}$–$Pro^{800}$) follows a disordered stretch of six residues and is related to the ligand-binding LRR domain by an ∼30° tilt out of plane in the direction of the ascending flank and by a ∼30-Å translation perpendicular to the β-sheet. It consists of a conventional *N*-cap ($Lys^{630}$–$Pro^{668}$; LRRNT2), three LRRs (LRR18–LRR20), and a classical C-terminal capping motif ($Trp^{743}$–$Pro^{800}$; LRRCT2). The N-terminal β-hairpin of LRRNT2 forms a short antiparallel β-sheet with the final LRRCT1 β-strand ($Ile^{617}$–$Pro^{619}$) of the large N-terminal domain.

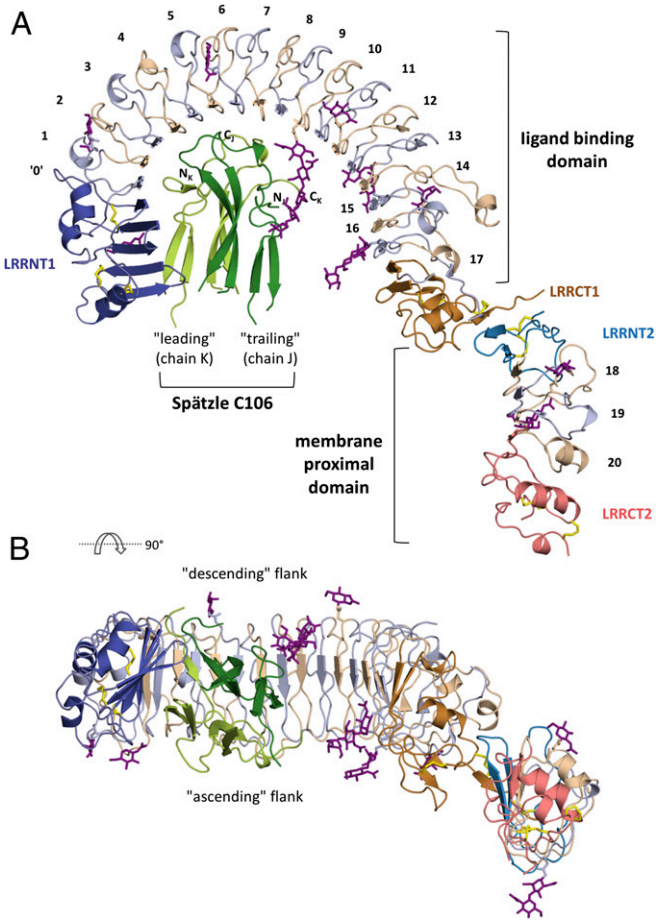


**Fig. 1.** Overall structure of the Toll-Spätzle complex. (*A*) The Toll ligand-binding domain is composed of three modules: an N-terminal capping motif (LRRNT1, dark blue), LRR motifs 0–17 (alternately colored light blue and beige), and a C-terminal capping motif (LRRCT1, orange). The membrane proximal domain consists of an N-terminal capping motif (LRRNT2, blue), LRRs 18–20 (color-coded as in the ligand-binding domain), and a C-terminal capping motif (LRRCT2, bright red). Intramolecular Toll ECD disulfide bonds are shown as yellow sticks; oligosaccharides, as purple sticks. The Spätzle cystine knot domain C106 is shown in light green (leading protomer, chain K) and dark green (trailing protomer, chain J). The N- and C-termini of each protomer are marked. (*B*) View rotated 90° about a horizontal axis, looking down the local two-fold axis of the Spätzle C106 dimer (color-coded as in *A*). Note the relationship between the Toll ligand-binding and membrane-proximal domains. N-linked glycosylation sites are seen on both the ascending (C-terminal to LRR β-strands) and descending (N-terminal to LRR β-strands) flanks of the horseshoe-shaped Toll ECD.

The N-linked glycosylation sites of the Toll ECD [of type $GlcNAc_2$-$Man_5$ (33)] are distributed on both the descending ($N^{175}$, $N^{270}$, and $N^{391}$) and ascending ($N^{80}$, $N^{140}$, $N^{235}$, $N^{482}$, and $N^{508}$) flanks as well as on the concave surface ($N^{346}$ and $N^{528}$) of the ligand-binding LRR domain, and on both the ascending ($N^{654}$ and $N^{703}$) and descending ($N^{715}$) flanks of the membrane-proximal domain (Fig. 1 and *SI Appendix*, Fig. S2 and Table S1). Of note, $N^{275}$, which is predicted to be a glycosylation site (35), is devoid of sugars and is intimately involved in ECD–Spätzle interaction (see below).

**Interaction with Spätzle.** Both protomers of the covalently linked Spätzle dimer interact with the concave surface of Toll (Fig. 2). The local twofold axis of the dimeric ligand is almost perpendicular to the tangent plane of the LRR semicircle, intersecting with the latter at $Ser^{277}$ at the C-terminal end of the LRR7

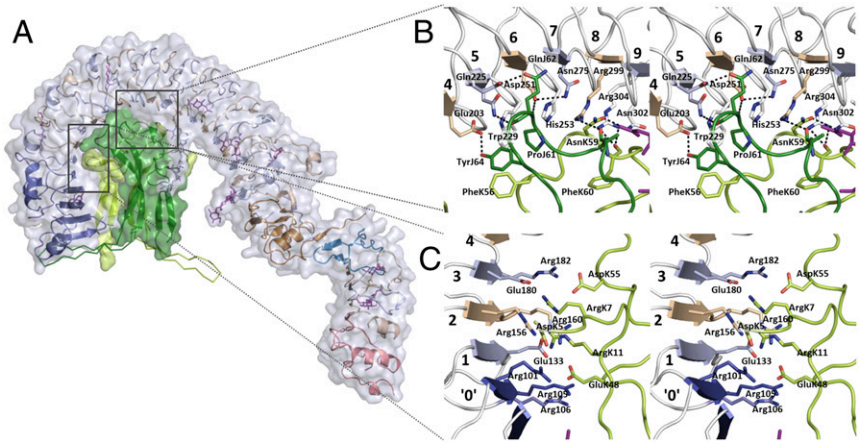


**Fig. 2.** Details of the Toll–Spätzle interaction. (*A*) Solvent-accessible surface of the complex (Toll, white surface; Spätzle protomers, green surfaces; secondary structure elements color-coded as in Fig. 1). The green Cα-backbone traces protruding from the surface indicate the positions of the Spätzle wings seen in the crystal structure of free Spätzle C106 (27) after superimposition on Toll-bound Spätzle. The trailing protomer wing (chain J, light green) would clash with residues of Toll LRRNT1. (*B*) Stereoview of binding interface LRR4–LRR9, demonstrating extensive hydrogen bonding and van der Waals contacts between Toll and both Spätzle protomers. Color-coding is as in Fig. 1, and dotted lines represent intermolecular hydrogen bonds between Toll and the Spätzle cystine knot doman. (*C*) Stereoview of the largely electrostatic binding interface involving Toll LRRs 0-3 and residues of the leading Spätzle protomer.

β-strand (Fig. 1*B*), so that steric restraints preclude binding of a second Toll ECD making use of the Spätzle symmetry. The morphogen-binding site is decidedly asymmetric; the "leading" Spätzle protomer (chain K) buries a surface area of 1,100 Å$^2$, whereas the "trailing" chain J buries 770 Å$^2$. The interface can be divided into three regions, the first of which is an elaborate electrostatic network involving residues Asp$^{K5}$, Arg$^{K7}$, Arg$^{K11}$, Glu$^{K48}$, and Asp$^{K55}$ of the leading Spätzle protomer and Toll residues Asp$^{100}$, Arg$^{101}$, Arg$^{105}$, Arg$^{106}$, Glu$^{133}$, Arg$^{154}$, Arg$^{160}$, Glu$^{180}$, and Arg$^{182}$ from LRR0–LRR3 (Fig. 2*C*). The position of conserved Arg$^{182}$ marks a one-residue LRR deletion in LRR3 (*SI Appendix*, Fig. S1) that likewise is highly conserved among insects (*SI Appendix*, Fig. S3).

The second region is formed by the base of the Spätzle dimer, which buries a substantive area of LRR4–LRR9, with ligand residues Phe$^{K56}$–Tyr$^{K64}$ and Ala$^{J57}$–Tyr$^{J64}$ arranged antiparallel to one another and almost perpendicular to the LRR β-strands (Fig. 2*B*). Extensive hydrogen bonding and van der Waals contacts between the ligand and receptor are observed in this region, centered on Spätzle residues Phe$^{K56}$, Asn$^{K59}$, Pro$^{J61}$, Gln$^{J62}$, and absolutely conserved Tyr$^{J64}$ (*SI Appendix*, Fig. S4) and on Toll residues Glu$^{203}$, Gln$^{225}$, Asn$^{227}$, Trp$^{229}$, Asp$^{251}$, His$^{253}$, Ser$^{277}$, Arg$^{299}$, and Asn$^{302}$. The fully buried Asn$^{275}$, held in place by the side chain of Asp$^{251}$, appears to stabilize the tautomer of His$^{253}$, allowing presentation of a hydrophobic surface to the side chain of Pro$^{J61}$. The guanidinium moieties of Arg$^{299}$ and Arg$^{304}$ each contribute hydrogen bonds to the main chain carbonyl groups of Ala$^{J58}$ and Phe$^{K60}$ in Spätzle. Arg$^{304}$ is found in an enlarged concave surface caused by a two-residue LRR insertion in LRR8 (*SI Appendix*, Fig. S1). This insertion disrupts the LRR Asn ladder backbone, as reflected by a narrower loop structure in the subsequent 23-aa LRR9 and a main chain hydrogen bond between Leu$^{307}$ and Ala$^{328}$.

Further contacts occur between the remaining residues of the Spätzle segments Phe$^{K56}$–Tyr$^{K64}$ and Ala$^{J57}$–Tyr$^{J64}$ to Ser$^{230}$, Asp$^{254}$, Ala$^{278}$, Met$^{301}$, Arg$^{325}$, and Arg$^{327}$. Finally, a smaller number of interactions are seen between one flank of Spätzle (residues Ile$^{K13}$, Lys$^{K15}$, Leu$^{K16}$, Gln$^{K43}$, Tyr$^{J72}$, Thr$^{J73}$, Gln$^{J74}$, and Lys$^{J85}$) and the N-terminal capping domain (Ile$^{48}$, Met$^{49}$, Glu$^{53}$, Arg$^{66}$, and Met$^{78}$), although the density is weak, and the model displays relatively high B-factors in this region (*SI Appendix*, Fig. S5).

**Removal of Only One Spätzle Prodomain Is Sufficient for Toll Binding, but Not for Dimerization.** Given that the prodomains in unprocessed proSpätzle are known to block binding to Toll (31, 33, 34), the question arises as to whether receptor binding requires removal of both prodomains. Refolding a 1:1 mixture of proSpätzle WT and a variant in which the Easter maturation site VSSR|VG is replaced by VSSAVG (28) allows isolation of a disulfide-linked chimeric heterodimer in which only one protomer contains the Easter cleavage site (*SI Appendix*, Fig. S6). After Easter treatment and removal of the single noncovalently bound prodomain, "half-cleaved" Spätzle (hcSpätzle) binds to the ECD with an affinity comparable to that for the C106–ECD interaction (Fig. 3 and *SI Appendix*, Table S2), providing support for the nonsymmetric binding mode observed in the crystals.

We and others (27, 34, 40, 41) have proposed that the prodomains cover a hydrophobic flank of the dimer between the base of the molecule and the wing. Assuming that the (covalently linked) uncleaved prodomain in hcSpätzle is not influenced by removal of the opposing prodomain, geometric constraints demand that the leading protomer juxtaposing LRRNT1 must be in the processed form, and that the remaining prodomain of hcSpätzle must be attached to the otherwise exposed trailing protomer.

Sedimentation velocity measurements demonstrate that hcSpätzle forms a 1:1 (monomeric) complex with the Toll ECD (Fig. 3*C*). In contrast, C106 forms a complex of dimeric stoichiometry (two ECDs–two cystine knot dimers, 2:2) under the same conditions, together with a minor fraction of a monomeric (1:1) species. The marginal (~ 2 kcal·mol$^{-1}$) difference in enthalpic contribution to ligand binding from the dimeric C106-Toll and monomeric hcSpätzle-Toll complexes (*SI Appendix*, Table S2) indicates that the C106-mediated Toll dimerization process is most likely entropy-driven. It should be noted, however, that a detailed interpretation of the calorimetric data are complicated by the presence of preformed Toll dimers (32, 34). Sedimentation velocity measurements (*SI Appendix*, Fig. S7) confirm the presence of a low-affinity dimer of the Toll ECD in the absence of ligand with a $K_d \geq 6$ μM; thus, a significant Toll$_2$ dimer fraction must exist at the concentrations needed for isothermal calorimetric experiments.

## Discussion

**Conformational Changes on Morphogen–Receptor Binding.** Superimposition of the recently determined structure of a truncated N-terminal domain of Toll (LRRNT1-LRR1-LRR3 fused to a C-terminal variable lymphocyte receptor domain) (35) with the full-length ECD shows negligible structural differences between the free and bound receptors. However, a comparison of the leading and trailing Spätzle chains reveals significant conformational differences between the two protomers at the base of the molecule (residues Phe$^{56}$–Tyr$^{64}$) (*SI Appendix*, Fig. S8), differences that are also observed, albeit to a lesser extent, in the "free" Spätzle structure (27). Further deviations from twofold symmetry are seen for Spätzle residues Glu$^{48}$–Gln$^{52}$; whereas these residues are solvent-exposed in the leading chain, Glu$^{J48}$, Gly$^{J49}$, Ala$^{J50}$, and Gln$^{J52}$ of the trailing protomer interact with

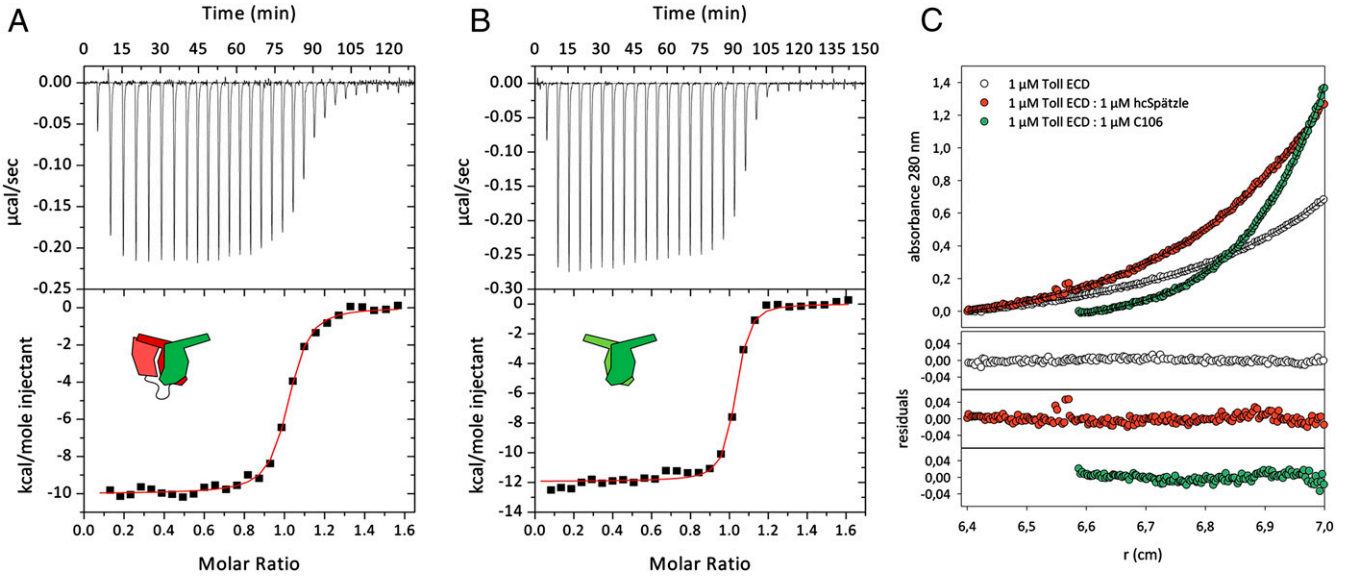


**Fig. 3.** The interaction of Toll ECD with hcSpätzle supports an asymmetric binding mode of Spätzle C106. (*A* and *B*) Isothermal titration calorimetry of 68 µM hcSpätzle (red/green pictogram) (*A*) or 68 µM Spätzle C106 (green) (*B*) to 9 µM Toll ECD reveals similar binding characteristics of the two ligand species. Measurements were performed at 25 °C in 20 mM Tris·HCl pH 7.4 and 150 mM NaCl. Data were corrected for the heat of dilution of the injectant before nonlinear regression of the binding curves using a model with one set of identical binding sites. (*C*) Sedimentation equilibrium measurements of Toll receptor indicate that the free glycosylated ECD (white; theoretical molecular mass, 110 kDa) is monomeric under the given conditions, with $M_r$ 100–115 kDa. A homogenous species with $M_r$ 260–270 kDa is found in complex with C106 (green), corresponding to a 2:2 ECD: C106 dimer stoichiometry (theoretical molecular weight 270 kDa) as described previously (34), whereas in complex with hcSpätzle, $M_r$ is 150 kDa (1:1 complex; theoretical value, 150 kDa). All data were obtained at 20 °C and 6,000 rpm (rotor An50Ti, Beckman Coulter) in 50 mM Tris·HCl pH 7.5, 150 mM NaCl, and 0.005% (wt/vol) Nonidet P-40. (*Lower*) Residuals of the corresponding nonlinear regressions.

the $Asn^{346}$ N-linked glycan residues $Man_3$ and $Man_4$. The most striking difference between Toll-bound Spätzle and the previously determined free structure (27) is the lack of ordered density in the morphogen wings, which include the highly conserved $Trp^{J/K29}$ (40, 41) (*SI Appendix*, Fig. S4). Superimposition of the two structures clearly shows that residues $Thr^{J76}$–$Gln^{J90}$ of the trailing monomer would clash with the Toll N-terminal domain (Fig. 2*A* and *SI Appendix*, Fig. S8*B*).

**Gain-of-Function *Toll* Mutants.** Quite apart from analogies to liganded TLRs, the majority of which crystallize as dimers (18, 19), EGF-Toll chimeras (22), along with strongly ventralized fly embryo phenotypes observed for a series of dominant gain-of-function alleles of mutant *Toll* genes, strongly suggest a role of dimerization in the activation of Toll signaling (*SI Appendix*, Fig. S9). Three of these mutants involve a G→A change in base 2 of one of the final three Cys codons in the LRRCT2 (6), causing the replacement of $Cys^{755}$, $Cys^{781}$, or $Cys^{799}$ by a bulky tyrosyl side chain. This would destabilize the membrane-proximal LRR domain and result in an unpaired cysteine, providing an explanation for the constitutive activity of these mutants through the formation of aberrant disulfide-linked oligomers (42).

The structure also provides a physicochemical rationale for a second class of dominant mutants that are viable only in the presence of WT *Toll* (6). The strongly ventralizing mutants $Toll^{84c}$ and $Toll^{5b}$ (caused by insertions of stop codons instead of $Gln^{464}$ and $Gln^{669}$, respectively) would result in secreted proteins C-terminally truncated at LRR14 or LRR18. Although these proteins should be able to bind Spätzle, we surmise that the lack of a C-terminal cap would result in strong destabilization, leading to a lower affinity for the ligand compared with an intact receptor. Thus, these secreted truncated proteins can redistribute Spätzle in the developing embryo away from the site of morphogen generation, leading to nonlocalized activation of Toll signaling and ventralization of the embryo, as originally suggested by the Anderson group (6). On the other hand, the weakly ventralizing $Toll^{2b}$ ($Gln^{614}$→stop) results in an almost entire N-terminal domain (LRRNT1 + LRR1-17 + LRRCT1). Spätzle can be redistributed, but owing to the structural integrity of the isolated domain, the ligand is rereleased in only minor quantities. Similar arguments can be made for the weakly ventralizing $Toll^{DB1}$. Because of a frameshift deletion (*SI Appendix*, Fig. S9), this protein also represents a nearly complete N-terminal domain, with the serendipitous replacement of $Cys^{597}$ by a cysteine residue created by the frameshift allowing the formation of a disulfide bridge with the otherwise unpaired $Cys^{565}$. Finally, the mutants $Toll^{DB2}$ and $Toll^{DB3}$ ($Trp^{723}$→stop and $Trp^{753}$→>stop, respectively), with their moderately truncated C-terminal domains, are likely to have intermediate stability, explaining their moderate ventralizing activity.

**Comparison with Mammalian TLRs.** Despite the obvious functional and (presumably) evolutionary relationship to mammalian TLRs, the Toll–Spätzle interaction appears to be fundamentally different (*SI Appendix*, Fig. S10). For TLRs, ligands bind to the ascending flank of the receptor, which accordingly is largely devoid of glycosylation. In all of the cases studied so far, TLR–ligand binding—in several cases effected by (pseudo) twofold ligand symmetries, such as for TLR3-dsRNA (43) or TLR5-flagellin (21)—has resulted in lateral dimerization of the receptor, leading to an intracellular platform for further signaling complexes. In contrast, Spätzle C106 binds to the concave surface of Toll, and as such is much more akin to most other LRR-protein ligand complexes first seen in the RNAseA-RNase inhibitor complex (44). Moreover, the twofold symmetry of the Spätzle ligand does not and cannot contribute to receptor dimerization. Conceptually, the Toll–Spätzle interaction resembles the follicle-stimulating hormone (FSH)–FSH receptor interaction (45, 46), in which the cystine knot head-to-tail heterodimer FSH binds to the LRR ECD of the G protein-coupled receptor FSH receptor

with its pseudo-twofold axis intersecting and perpendicular to the LRR concave surface.

**Toll Oligomerization.** A low-resolution cryoEM single-particle reconstruction of a dimeric Toll-Spätzle assembly has been reported (34) in which Spätzle is proposed to bind to the tip of LRRNT1, which is not possible based on the recent characterization of a truncated Toll N-terminal domain (35). Comparing the form and dimensions of the full-length ECD structure presented here with the cryoEM map suggests that the particles used for the reconstruction correspond to the preformed Toll dimer, with the volume assigned previously to Spätzle corresponding to the Toll ECD C-terminal membrane-proximal domain (*SI Appendix*, Fig. S11).

Correlating the present structure with the density suggests that the LRRNT1s of two Toll ECDs come head-to-head in a preformed dimer, with the ascending flanks of LRR12 to LRRCT1 of each monomer approaching one another and the membrane-proximal domain C-termini of the two receptors some 160 Å apart. Assuming such an arrangement, the superimposition of the 1:1 Toll-Spätzle complex on each ECD suggests two consequences of Spätzle binding: interference of head-to-head LRRNT1 contacts by the leading wings of a single C106 dimer and lateral steric clashes of the two bound cystine knot dimers.

How can the present structural data be reconciled with the initiation of Toll signaling? The preformed Toll dimer appears to represent an inactive state of the receptor, similar to the situation described recently for the unliganded TLR8 (20), kept in place by stabilizing interactions between their juxtaposing N-terminal domains. We suggest that binding of Spätzle disrupts this complex, leading to a rearrangement of the dimer that allows a close approach of the C-terminal membrane proximal domains to initiate Toll signaling (*SI Appendix*, Figs. S11 and S12), although whether this is a 1:2 or a 2:2 complex remains an open question.

The absence of an oligomeric signaling complex in our crystals is clearly perplexing, however. Although it is well established that weak protein–protein interactions in solution can be enhanced considerably on 2D confinement within a membrane (47) [e.g., the 2:2 TLR5:Flic dimer responsible for flagellin recognition was not detectable using gel filtration and revealed a $K_d$ of only ~5 mM using analytical ultracentrifugation (21)], all TLR-agonist complexes studied to date show dimer formation at the high protein concentrations prevalent during crystallization (18–21). Although it is possible that the 1:1 Toll-Spätzle dimer observed here may represent a presignaling complex, the situation is reminiscent of the TLR4-MD2 complex, in which initial elucidation of the structure revealed a monomeric complex (38). The dimeric signaling complex was obtained only in the presence of a suitable lipopolysaccharide ligand bound to MD2 in the complex (48). By analogy, Toll-Spätzle signaling may be mediated by an as-yet unidentified coligand.

We noted previously (27) that the Spätzle C106 dimer has a deep internal cavity (the entry of which is on the twofold axis near to the wings and the otherwise disordered $Trp^{29}$ loop) that can accept hydrophobic ligands in a manner similar to MD2. It was recently demonstrated that the structurally related NGF molecules from cobra venom and mouse are able to bind lipids in a corresponding region, and that lipid binding modulates the structure and activity of neurotrophin (49); thus, it is not inconceivable that binding of an unknown coligand to Spätzle [which is not likely to be present in our *Escherichia coli* expression and in vitro refolding system but might be available in insect cell culture- produced protein (31, 32, 34)] organizes residues of the Spätzle wings to facilitate receptor dimerization and signaling (*SI Appendix*, Fig. S12*A*).

It is also possible that the prodomains of Easter-generated Spätzle*, which are displaced on binding to Toll (29), influence complex formation. The finding that hcSpätzle interacts with the Toll ECD in an identical fashion to C106 implies that on binding to Toll, only one prodomain [which binds with only micromolar affinity to C106 (28)] of Easter-cleaved Spätzle* needs to be displaced to form the complex observed here. In turn, the remaining (cleaved yet associated) prodomain in an intermediate 1:1:1 prodomain-C106 dimer-Toll complex could support receptor dimerization and thus signaling (*SI Appendix*, Fig. S12*B*).

Both of the foregoing activation scenarios raise the intriguing possibility that the signaling complex may dissociate even at high concentrations of the Spätzle ligand. A key feature of developmental pattern formation is the establishment of a morphogen gradient distribution to produce differentiated cell types in a location-dependent manner; however, at early stages of development, sharp gradients are difficult to achieve owing to more-or-less homogeneous distributions of both morphogens and cell types (reviewed in refs. 13, 50). Given that overexpression of proSpätzle in vivo leads to axis duplication (51), it has been suggested that the Spätzle prodomain might act as a diffusible inhibitor of upstream Toll activation processes, possibly via a shuttling mechanism (52); however, until now, how this inhibition might occur was not clear, particularly because the isolated prodomain has been shown to be unstructured (27–29) and N-terminal prodomain deletion mutants do not exhibit lateral domain expansion in vivo (53). Dissociation of dimeric signaling complexes at high Spätzle concentrations effected by, for instance, limitation of coligand or prodomain removal could provide an inbuilt shutdown mode to the Toll-Spätzle cascade, allowing establishment of a sharply defined zone of Toll signaling and hence differentiation. We note in passing that a 1:1 Toll-Spätzle complex has been reported in the presence of excess Spätzle (32), although we have not observed 2:2 dimer dissociation at high ligand concentrations.

Insects represent one of the most diverse animal classes, with annotations for more than 1 million species. Along with their contributions to biodiversity, many insect species are of economic importance, obvious examples being crop pollination by the honeybee *Apis mellifera* and silk production by the silkworm *Bombyx mori*. On the other hand, insect vectors carry a wide range of diseases, including malaria (*Anopheles gambiae*) and yellow fever (*Aedes aegypti*), and a number of insects are downright pests, such as the body louse *Pediculus corporis*. The genomes of these species each have at least one homolog to both Spätzle and Toll from *D. melanogaster*. Conservation of residue patterns in the two proteins identified as important for their interaction (*SI Appendix*, Fig. S3 and S4) points to common modes of receptor ligand recognition, similarities that extend to other arthropods, such as the water flea *Daphnia pulex* (a crustacean) and the hard tick *Ixodes scapularis* (an arachnid and the vector for Lyme disease). Differences in the details of these interactions raise the possibility of targeting the immune or developmental systems of specific insects, with a view toward either strengthening the immune response (e.g., of honey bees, the world population of which is currently under threat from the arachnid *Varroa destructor*) or the opposite (e.g., for malaria control). In addition to the surprising features revealed by this pioneering protein pair, there remains much to be gained from a greater understanding of the molecular details of invertebrate Toll-Spätzle recognition.

## Methods

**Expression, Complex Formation, and Crystallization.** The cloning, expression, purification, complex formation, and crystallization of the deglycosylated complex have been described previously (33). Further details are provided in *SI Appendix, Methods*.

**Structure Determination.** Phases for the deglycosylated complex were determined using multiple-wavelength anomalous dispersion with a single samarium derivative, and the final model of the complex was refined against native data to 2.2-Å resolution. No interpretable electron density was observed for Toll residues $His^{544}$–$Tyr^{549}$, $Asp^{624}$–$Arg^{629}$, $Arg^{739}$–$Lys^{742}$, $Glu^{785}$–$Met^{786}$, or $Ala^{801}$–$His^{810}$ (the latter including the C-terminal $His_6$ tag); for

$Val^{J1}$–$Ser^{J4}$ of the trailing Spätzle protomer; or for the Spätzle wings ($Tyr^{K18}$–$Gln^{K39}$/$Gln^{J75}$–$Phe^{J93}$ and $Gln^{K75}$–$Phe^{K93}$/$Tyr^{J18}$–$Gln^{J40}$). Similarly, the final residues of the Spätzle C-terminal $His_6$ tags are not defined, although density is present for residues $Gly^{J106}$–$His^{J110}$ and $Gly^{K106}$–$Leu^{K107}$, which are included in the model. The completed deglycosylated structure was used to solve the nearly isomorphous crystals of the fully glycosylated complex, which after dehydration diffracted to 3.5-Å resolution. Crystallographic details are provided in *SI Appendix, Methods*.

**Preparation of Spätzle* and hcSpätzle and Analysis of Their Interactions with Toll.** All Spätzle variants were derived from isoform Spz11.7, expressed as inclusion bodies of proSpätzle in *E. coli*, and refolded as described previously (27). The half-cleaved hcSpätzle variant was generated by refolding an equimolar (denatured) mixture of proSpätzle and a noncleavable variant [in which the Easter maturation site VSSR-VGG is replaced by VSSA-VGG (28)] lacking the C-terminal $His_6$ tag. Activation cleavage was effected using recombinant $Easter^{Xa}$ (28) to yield hcSpätzle* (or Spätzle* from homodimeric refolded proSpätzle), and the noncovalently bound prodomains were removed under mildly denaturing conditions. The interactions of the thus-generated hcSpätzle and C106 with the Toll ECD were analyzed using isothermal calorimetry and analytical ultracentrifugation, as described in *SI Appendix, Methods*.

**Note Added in Proof.** While this paper was under review, Lewis et al. (54) reported the structure of Spätzle in complex with a truncated Toll ectodomain consistent with the data presented here.

**ACKNOWLEDGMENTS.** We thank Uwe Müller [Free University Berlin at the Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (BESSY)] for providing synchrotron beam time and Christian Schwieger (Martin-Luther-Universität Halle-Wittenberg) for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich SFB610, "Protein-Zustände mit zellbiologischer und medizinischer Relevanz."

1. Anderson KV, Nüsslein-Volhard C (1984) Information for the dorsal–ventral pattern of the *Drosophila* embryo is stored as maternal mRNA. *Nature* 311(5983):223–227.
2. Anderson KV, Bokla L, Nüsslein-Volhard C (1985) Establishment of dorsal-ventral polarity in the *Drosophila* embryo: The induction of polarity by the Toll gene product. *Cell* 42(3):791–798.
3. Anderson KV, Jürgens G, Nüsslein-Volhard C (1985) Establishment of dorsal-ventral polarity in the *Drosophila* embryo: Genetic studies on the role of the Toll gene product. *Cell* 42(3):779–789.
4. Hashimoto C, Hudson KL, Anderson KV (1988) The Toll gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell* 52(2):269–279.
5. Imler JL, Hoffmann JA (2001) Toll receptors in innate immunity. *Trends Cell Biol* 11(7): 304–311.
6. Schneider DS, Hudson KL, Lin TY, Anderson KV (1991) Dominant and recessive mutations define functional domains of Toll, a transmembrane protein required for dorsal-ventral polarity in the *Drosophila* embryo. *Genes Dev* 5(5):797–807.
7. Gay NJ, Keith FJ (1991) *Drosophila* Toll and IL-1 receptor. *Nature* 351(6325):355–356.
8. Schneider DS, Jin YS, Morisato D, Anderson KV (1994) A processed form of the Spätzle protein defines dorsal-ventral polarity in the *Drosophila* embryo. *Development* 120(5): 1243–1250.
9. Stein D, Roth S, Vogelsang E, Nüsslein-Volhard C (1991) The polarity of the dorsoventral axis in the *Drosophila* embryo is defined by an extracellular signal. *Cell* 65(5): 725–735.
10. DeLotto Y, DeLotto R (1998) Proteolytic processing of the *Drosophila* Spätzle protein by easter generates a dimeric NGF-like molecule with ventralising activity. *Mech Dev* 72(1-2):141–148.
11. LeMosy EK, Hong CC, Hashimoto C (1999) Signal transduction by a protease cascade. *Trends Cell Biol* 9(3):102–107.
12. Morisato D, Anderson KV (1995) Signaling pathways that establish the dorsal-ventral pattern of the *Drosophila* embryo. *Annu Rev Genet* 29:371–399.
13. Moussian B, Roth S (2005) Dorsoventral axis formation in the *Drosophila* embryo: Shaping and transducing a morphogen gradient. *Curr Biol* 15(21):R887–R899.
14. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* 86(6):973–983.
15. Medzhitov R, Preston-Hurlburt P, Janeway CA, Jr. (1997) A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* 388(6640): 394–397.
16. Poltorak A, et al. (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene. *Science* 282(5396):2085–2088.
17. Gay NJ, Gangloff M (2007) Structure and function of Toll receptors and their ligands. *Annu Rev Biochem* 76:141–165.
18. Botos I, Segal DM, Davies DR (2011) The structural biology of Toll-like receptors. *Structure* 19(4):447–459.
19. Kang JY, Lee JO (2011) Structural biology of the Toll-like receptor family. *Annu Rev Biochem* 80:917–941.
20. Tanji H, Ohto U, Shibata T, Miyake K, Shimizu T (2013) Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science* 339(6126):1426–1429.
21. Yoon SI, et al. (2012) Structural basis of TLR5-flagellin recognition and signaling. *Science* 335(6070):859–864.
22. Sun H, Towb P, Chiem DN, Foster BA, Wasserman SA (2004) Regulated assembly of the Toll signaling complex drives *Drosophila* dorsoventral patterning. *EMBO J* 23(1): 100–110.
23. Tauszig S, Jouanguy E, Hoffmann JA, Imler JL (2000) Toll-related receptors and the control of antimicrobial peptide expression in *Drosophila*. *Proc Natl Acad Sci USA* 97(19):10520–10525.
24. McIlroy G, et al. (2013) Toll-6 and Toll-7 function as neurotrophin receptors in the *Drosophila melanogaster* CNS. *Nat Neurosci* 16(9):1248–1256.
25. Jang IH, et al. (2006) A Spätzle-processing enzyme required for toll signaling activation in *Drosophila* innate immunity. *Dev Cell* 10(1):45–55.
26. DeLotto Y, Smith C, DeLotto R (2001) Multiple isoforms of the *Drosophila* Spätzle protein are encoded by alternatively spliced maternal mRNAs in the precellular blastoderm embryo. *Mol Gen Genet* 264(5):643–652.
27. Hoffmann A, et al. (2008) Biophysical characterization of refolded *Drosophila* Spätzle, a cystine knot protein, reveals distinct properties of three isoforms. *J Biol Chem* 283(47):32598–32609.
28. Ursel C, et al. (2013) In vitro maturation of *Drosophila* melanogaster Spätzle protein with refolded Easter reveals a novel cleavage site within the prodomain. *Biol Chem* 394(8):1069–1075.
29. Weber ANR, et al. (2007) Role of the Spatzle Pro-domain in the generation of an active toll receptor ligand. *J Biol Chem* 282(18):13522–13531.
30. Mizuguchi K, Parker JS, Blundell TL, Gay NJ (1998) Getting knotted: A model for the structure and activation of Spätzle. *Trends Biochem Sci* 23(7):239–242.
31. Weber ANR, et al. (2003) Binding of the *Drosophila* cytokine Spätzle to Toll is direct and establishes signaling. *Nat Immunol* 4(8):794–800.
32. Weber ANR, Moncrieffe MC, Gangloff M, Imler JL, Gay NJ (2005) Ligand–receptor and receptor–receptor interactions act in concert to activate signaling in the *Drosophila* toll pathway. *J Biol Chem* 280(24):22793–22799.
33. Stelter M, et al. (2013) High level expression of the *Drosophila* Toll receptor ectodomain and crystallization of its complex with the morphogen Spätzle. *Biol Chem* 394(8):1091–1096.
34. Gangloff M, et al. (2008) Structural insight into the mechanism of activation of the Toll receptor by the dimeric ligand Spätzle. *J Biol Chem* 283(21):14629–14635.
35. Gangloff M, Arnot CJ, Lewis M, Gay NJ (2013) Functional insights from the crystal structure of the N-terminal domain of the prototypical Toll receptor. *Structure* 21(1): 143–153.
36. Bella J, Hindle KL, McEwan PA, Lovell SC (2008) The leucine-rich repeat structure. *Cell Mol Life Sci* 65(15):2307–2333.
37. Buchanan SG, Gay NJ (1996) Structural and functional diversity in the leucine-rich repeat family of proteins. *Prog Biophys Mol Biol* 65(1-2):1–44.
38. Kim HM, et al. (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* 130(5):906–917.
39. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* 11(6):725–732.
40. Arnot CJ, Gay NJ, Gangloff M (2010) Molecular mechanism that induces activation of Spätzle, the ligand for the *Drosophila* Toll receptor. *J Biol Chem* 285(25):19502–19509.
41. Wang Y, Zhu S (2009) Evolutionary and functional epitopes of the Spätzle protein: New insights into activation of the Toll receptor. *Cell Mol Life Sci* 66(9):1595–1602.
42. Hu XD, Yagi Y, Tanji T, Zhou SL, Ip YT (2004) Multimerization and interaction of Toll and Spätzle in *Drosophila*. *Proc Natl Acad Sci USA* 101(25):9369–9374.
43. Liu L, et al. (2008) Structural basis of Toll-like receptor 3 signaling with double-stranded RNA. *Science* 320(5874):379–381.
44. Kobe B, Deisenhofer J (1995) A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* 374(6518):183–186.
45. Fan QR, Hendrickson WA (2005) Structure of human follicle-stimulating hormone in complex with its receptor. *Nature* 433(7023):269–277.
46. Jiang X, et al. (2012) Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. *Proc Natl Acad Sci USA* 109(31):12491–12496.
47. Gavutis M, Jaks E, Lamken P, Piehler J (2006) Determination of the two-dimensional interaction rate constants of a cytokine receptor complex. *Biophys J* 90(9):3345–3355.
48. Park BS, et al. (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* 458(7242):1191–1195.
49. Tong Q, et al. (2012) Structural and functional insights into lipid-bound nerve growth factors. *FASEB J* 26(9):3811–3821.
50. Shilo BZ, Haskel-Ittah M, Ben-Zvi D, Schejter ED, Barkai N (2013) Creating gradients by morphogen shuttling. *Trends Genet* 29(6):339–347.
51. Morisato D (2001) Spätzle regulates the shape of the dorsal gradient in the *Drosophila* embryo. *Development* 128(12):2309–2319.
52. Haskel-Ittah M, et al. (2012) Self-organized shuttling: Generating sharp dorsoventral polarity in the early *Drosophila* embryo. *Cell* 150(5):1016–1028.
53. Morisato D, Anderson KV (1994) The *spätzle* gene encodes a component of the extracellular signaling pathway establishing the dorsal-ventral pattern of the *Drosophila* embryo. *Cell* 76(4):677–688.
54. Lewis M, et al. (2013) Cytokine Spätzle binds to the *Drosophila* immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. *Proc Natl Acad Sci USA* 110(51):20461–20466.